REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
31-05-2011	Final	Apr 2008--May 2011

4. TITLE AND SUBTITLE

Solving Boltzmann and Fokker-Planck equations using sparse representation

5a. CONTRACT NUMBER

5b. GRANT NUMBER
FA9550-08-1-0416

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Jie Shen and Weinan E

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Purdue University
West Lafayette, IN 47907

and

Princeton University
Princeton, NJ 08544

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

AF Office of Scientific Research
875 N. Randolph St. room 3112
Arlington, VA 22203

10. SPONSOR/MONITOR'S ACRONYM(S)

AFOSR

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AFRL-OSR-VA-TR-2012-0877

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release

13. SUPPLEMENTARY NOTES

14. ABSTRACT

A major issue in modeling and computation is how to handle high dimen- sional problems. We can divide these high dimensional problems into two classes: moderately high dimensional problems or very high dimensional prob- lems. In the former class, we have problems such as the Boltzmann equation and Fokker-Planck equation, whose dimensionality is moderately high but are amendable to sparse grid based methods. In the latter class, we have problems such as exploration of the configuration space of a large molecule. These prob- lems often involve hundreds of thousands of dimensions, and methods based on fixed grids are far from being adequate. We developed various techniques in handling these problems using the hyperbolic cross/sparse representation for the former class, and adaptive sampling for the latter. These developments are aimed at providing a solid foundation for efficient and reliable numerical simulations of Boltzmann and Fokker-Planck equations.

Besides the work presented in this report, a number of other related publi- cations by the PIs were also partially supported by this grant.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
					Jie Shen
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER** *(include area code)*
UU	UU	UU			7654941923

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

SOLVING BOLTZMANN AND FOKKER-PLANCK EQUATIONS USING SPARSE REPRESENTATION

AFOSR FA9550-08-1-0416

Jie Shen
Department of Mathematics, Purdue University

Weinan E
Department of Mathematics, Princeton University

Abstract

A major issue in modeling and computation is how to handle high dimensional problems. We can divide these high dimensional problems into two classes: moderately high dimensional problems or very high dimensional problems. In the former class, we have problems such as the Boltzmann equation and Fokker-Planck equation, whose dimensionality is moderately high but are amendable to sparse grid based methods. In the latter class, we have problems such as exploration of the configuration space of a large molecule. These problems often involve hundreds of thousands of dimensions, and methods based on fixed grids are far from being adequate. We developed various techniques in handling these problems using the hyperbolic cross/sparse representation for the former class, and adaptive sampling for the latter. These developments are aimed at providing a solid foundation for efficient and reliable numerical simulations of Boltzmann and Fokker-Planck equations.

Besides the work presented in this report, a number of other related publications by the PIs were also partially supported by this grant.

1 Sparse spectral methods for solving high-dimensional partial differential equations

How to numerically solve Boltzmann equations and Fokker-Planck equations are a major challenge in computational science. Traditional numerical approaches are hindered by the high dimensionality of these equations. An effective strategy for dealing with moderately high dimensional problems, as in the case of Boltzmann equations and Fokker-Planck equations, is to use hyperbolic cross/sparse grid based approximations.

Our first goal is to establish some fundamental approximation results for the hyperbolic cross approximations. The second goal is to develop efficient sparse spectral methods for numerical integrations on high-dimensional infinite domains and for solving a class of high-dimensional PDEs. The third goal is to develop efficient and accurate sparse spectralmethods to solve the Boltzmann and Fokker-Planck equations. We have essentially achieved the first two goals above, and have made substantial progress towards the third goal. Our progress in these directions are summarized below.

1.1 Error analysis for hyperbolic cross approximations

We studied the hyperbolic cross approximations based on Jacobi polynomials for bounded domains and Hermite/Laguerre functions on unbounded domains, and established optimal error estimates in proper anisotropic weighted Korobov spaces for both the regular hyperbolic cross approximations and optimized hyperbolic cross approximations. These results (cf. [1]) are of fundamental importance and are proved systematically with a uniform approach that can be used to study the hyperbolic cross approximations by other orthogonal systems.

1.2 Efficient sparse spectral methods for bounded and unbounded domains

In the case, of bounded domains, by using the generalized Jacobi polynomials $J_n^{-1,-1}(x)$ as basis functions, we showed that the linear system resulting from the hyperbolic cross approximation is sparse. However, one can no longer apply the usual technique (for the full grid) of matrix diagonalization/decomposition to efficiently solve this linear system. On the other hand, the integrals involving the forcing function f has to be approximated by a suitable quadrature or the function f has to be replaced by a suitable interpolation $I_N f$. We have developed a quasi-optimal sparse Chebyshev-Legendre method in which the Legendre formulation is used in order to have a sparse system matrix, while the sparse grid based on the Chebyshev-Gauss-Lobatto points is used for interpolation/integration. This new method makes it possible to solve moderately high dimensional problems which are otherwise out of reach by the usual full grid methods (cf. [2]).

However, in the case of unbounded domains, the classical quadrature rules based on Hermite and Laguerre functions are not nested, so the numbers of nodes in the Smolyak's sparse grid based on these quadrature rules grows significantly faster than the sparse grid based on a nested quadrature rule. To overcome this difficulty, we considered a mapped approach. Namely, we construct a family of mappings which map $(-1, 1)$ to $(-\infty, +\infty)$ or $(0, +\infty)$, and then use the mapped sparse grid based on the Chebyshev Gauss-Lobatto quadrature. The parameter in this family of mappings can be used to fine tune the performance with respect to the regularity and asymptotic property of the underlying function. We have developed efficient implementations by using these mapped Chebyshev sparse grids for numerical integration

as well as for solving model elliptic equations, and investigated their convergence rate and performance in comparison with the usual approach (cf. [3]).

In a related work [4], we developed an efficient stochastic Galerkin method for random diffusion equations.

1.3 Approximation of the Fokker-Planck Equation of FENE dumbbell model

A necessary step towards solving the full Navier-Stokes Fokker-Planck equations is to develop an efficient and accurate method for solving the Fokker-Planck equation. To fix the idea, we took the Fokker-Planck equation for FENE dumbbell model as an example. The FENE dumbbell model is a well-known coarse grain model for dilute polymer solutions and has been studied extensively in recent years. However, the existing mathematical formulation for the FENE dumbbell model assumes the initial condition decays sufficiently fast near the boundaries, and limits the scope of its applicability.

We introduced a new weighted weak formulation which allows the largest possible set of initial conditions, and proved its well-posedness in weighted Sobolev spaces. Moreover, we showed that the solution of our weighted weak formulation enjoys a boundary smoothing property. This was the first result of such kind for the FENE dumbbell model and we believe that the weighted weak formulation that we proposed is a natural formulation for the Fokker-Planck equation of the FENE dumbbell model.

We also constructed simple, yet efficient semi-implicit time-discretization schemes and proved that they are unconditionally stable. These semi-implicit schemes allow us to reduce 2-D and 3-D problems into a sequence of 1-D problems which can then be solved by tailored Jacobi spectral-Galerkin algorithms which enjoy the optimal computational complexity, conserve the volume naturally, and provide accurate approximation to higher-order moments of the distribution function. Our algorithms are orders of magnitude more efficient than existing schemes. Thus, they provide a solid first step towards our ultimate goal of directly solving the coupled five- and six-dimensional Navier-Stokes Fokker-Planck equations. These results are presented in [5] and [6].

1.4 Sparse spectral methods for electronic Schrödinger equation

2 Exploring parameter space using sparse representation/adaptive sampling

2.1 Calibrating and improving empirical inter-atomic potentials

One of the most important issues in atomistic modeling is to determine the inter-atomic potential. The process often goes as follows: one first specifies a form of the potential, based on experience or known facts which often involve many undetermined parameters; one then determines values of these parameters using experimental data or data obtained from first principle calculations. Needless to say, the success of such empirical potentials depends heavily on the assumed form.

Assessing the accuracy of such an empirical potential is a rather difficult task since the dimension of the configuration space of the atoms is usually very high. One question we have pursued is to assess the validity of the embedded atom model in the elastic regime. The embedded atom model is a very popular model for studying metallic systems. The space of elastic deformations is parameterized by the strain, which is a six-dimensional space. Using sparse representation in the six dimensional space, we were able to systematically explore the space of elastic deformations. It was found that the embedded atom model works quite well when the system is under shear or tension, but it works poorly when the system is under compression. In analogy with the modeling of exchange-correlation functionals in density functional theory and viewing the embedded atom model as the analog of the local density approximation (LDA), we have developed an analog of the generalized gradient approximation (GGA). Current results suggest that this improved model gives much better results in all regimes. The results are published in [7].

2.2 Sequential multiscale modeling

While sequential multiscale modeling has certain advantages, it if often believed that it is limited to the passage of a few parameters from microscale models to macroscale models. If the unknown component of the macroscale model is a function that depends on many variables, sequential multiscale modeling generally becomes ineffective. However, the power of sequential multiscale modeling can be greatly improved if sparse representations are used. For example, assume that we are modeling complex materials or complex fluids, for which the constitutive law for the stress is either a function of the strain or the rate of strain, which in three dimension is a function of six variables, then it becomes feasible to precompute the constitutive information using sparse representation. This has important consequences in multiscale modeling: In many cases, it is enough to use sequential multiscale modeling if we represent the constitutive information in a smart way. This is illustrated in [8].

2.3 The adaptive minimum action method

Sparse representation is only effective if the dimension of the problem is only moderately high. In computational science we often encounter problems whose dimension can be thousands or even larger. In this case, we have to use other strategies.

One strategy that we have pursued so far is adaptive sampling. The general procedure is as follows: we first identify a lower dimensional region (e.g. points, curves or surfaces) in the configuration spaces; we then sampling the probability distribution of our interest in the neighborhood of that region, and determine how the region moves in the iteration. Such an adaptive sampling procedure has been used with great success to gradient systems, i.e. systems that have an underlying energy landscape, with applications to many important problems in material science, chemistry and biology.

Our new interest is to study non-gradient systems.We will use transition pathways between stable states as an example.

Given two stable or metastable states, we would like to find the most probable transition path between the two states. For gradient systems, the string methods have been very successful. For non-gradient systems, we are exploring the minimum action method, which is more general. But for it to be effective, a high quality mesh along the path is required in order to resolve the path. This is very important since these paths are often very complicated.

We have developed an adaptive minimum action method. The basic idea comes from the moving mesh method. The objective is to find the optimal mesh using carefully chosen monitor functions, as the iteration proceeds. This is very simple, but it proves to be quite effective, see [9].

2.4 Application to the Kuramoto-Sivashinsky equation

As an application, we have used the adaptive minimum action method to study the dynamics of the Kuramoto-Sivashinsky equation in its configuration space. One first identifies a stable stationary solution and another traveling wave solution. By finding transition pathways between these two solutions, one further finds more stationary solutions. More importantly, one can further refine the whole procedure and find important objects on the separatrices between the basins of attractions of the different stable solutions. This procedure gives us a rich set of information about the very complicated configuration space of the Kuramoto-Sivashinsky equation. The results are documented in [10].

So far we have only studied the Kuramoto-Sivashinsky equation. But the methodology is fairly general. We intend to use this strategy to study the Navier-Stokes equations.

2.5 Subcritical instabilities

Many important physical systems exhibit subcritical bifurcations. One of the best known examples is laminar flow in a circular pipe, which is linearly stable for all

Reynolds number, yet it undergoes transition to more complicated and eventually turbulent flows when the Reynolds number is sufficiently large. In contrast to supercritical instabilities which are local phenomena in the configuration space and can be studied by analyzing linearized models, subcritical instabilities are related to the global behavior of the system under consideration.

It is clear that finite amplitude perturbations are needed in order to trigger subcritical instabilities. What is not clear, however, is how to turn this intuition into a set of tools with which one can analyze subcritical instabilities and make quantitative predictions. We have developed a quantitative tool for analyzing such instabilities, assuming that the instability is driven by a small amplitude noise. Our framework is based on the large deviation theory, and it gives specific criteria for determining the stability of a state under noisy perturbations, see [11]. We are now in the process of applying these results to the analysis of the instability for laminar flows.

3 Evaluation of the selected components of an inverse matrix

In many scientific applications, we need to calculate a subset of the entries of the inverse of a given matrix. A particularly important example is in the electronic structure analysis of materials using algorithms based on pole expansion [12] where the diagonal and sometimes sub-diagonals of the discrete Green's function or resolvent matrices are needed in order to compute the electron density. Other examples in which particular entries of the Green's functions are needed can also be found in the perturbation analysis of impurities by solving Dyson's equation in solid state physics, or the calculation of retarded and less-than Green's function in electronic transport. We will call this type of calculations selected inversion of a matrix. Our goal is to find a direct method to extract selected components, especially all the diagonal elements of the inverse of a given symmetric sparse matrix A without calculating the entire inverse matrix.

An obvious way to obtain selected components of A^{-1} is to compute A^{-1} first and then simply pull out the needed entries. The computational cost of such method is $\mathcal{O}(N^3)$ where N is the dimension of the system. This procedure already becomes prohibitively expensive for $N \sim 100,000$.

We have developed the selected inversion algorithm [13, 14, 15] to exploit the sparsity structure of A and obtain the selected components of the inverse matrix. Our algorithm is based on LDL^T factorization with necessary reordering strategy of the matrix A. The main result is that the inverse of A restricted to the non-zero pattern of L and L^T can be computed without seeking elements outside this pattern. Especially, the diagonal elements of the inverse matrix are inside this non-zero pattern. The number of non-zero elements in L is usually less than $\mathcal{O}(N^3)$, and the fast algorithm is obtained. To be more specific, the complexity of the selected inversion algorithm is $\mathcal{O}(N)$ for one dimensional system, $\mathcal{O}(N^{1.5})$ for two-dimensional system and $\mathcal{O}(N^2)$ for three-dimensional system.

The selected inversion algorithm is vastly suitable for the computation using different discretization techniques for the operator and different domain shapes. We have applied the selected inversion algorithm to evaluate the inverse of 2D Laplacian operator. The domain size is 65535×65535, which results in a matrix of size 4.3 billion. The selected inversion algorithm is able to compute the diagonal of the inverse of the 2D Laplacian operator within 25 minutes on $4,096$ processors.

Combined with the recently developed pole expansion technique, we have applied the selected inversion algorithm to Kohn-Sham density functional theory for metallic system, which is a well-known difficult problem in material science. We have computed the electronic structure of 2D quantum dot system, and compared the efficiency with the benchmark software OCTOPUS. For one self-consistent iteration step with 512 electrons, OCTOPUS costs 1091 sec, and selected inversion costs 9.76 sec. The algorithm exhibits significant advantage for 2D systems, and shows potential value for studying the electronic structure of 3D systems.

The selected inversion algorithm serves as an powerful alternative approach for large scale Kohn-Sham density functional theory calculation and ab initio molecular dynamics simulation with improved scaling property. The algorithm can be also applied to other related fields where selected components of inverse matrix is to be extracted, such as the non-equilibrium Green's functional approach for quantum transport computation.

Acknowledgment/Disclaimer
This work was sponsored (in part) by the Air Force Office of Scientific Research, USAF, under grant/contract number FA9550-08-1-0416. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Air Force Office of Scientific Research or the U.S. Government.

Personnel Supported During Duration of Grant
Weinan E, Professor, Princeton University
Jie Shen, Professor, Purdue University
Lin Lin, graduate student, Princeton University
Xiaoliang Wan, Postdoc, Princeton University
Haijun Yu, Postdoc, Purdue University
Xiang Zhou, graduate student, Princeton University

Publications:
[1]. Jie Shen and Li-Lian Wang, "Sparse spectral approximations of high dimensional problems based on hyperbolic cross", SIAM J. Numer. Anal., 48:1087-1109, 2010.
[2]. Jie Shen and Haijun Yu, "Efficient Spectral Sparse Grid Methods for High Dimensional Elliptic Problems", SIAM J. Sci. Comput. 32:3228-3250. 2010.
[3]. Jie Shen and Haijun Yu, "Mapped Chebyshev Sparse Grid Approximation on Unbounded Domain", submited to SIAM J. Sci. Comput.
[4]. Dongbin Xiu and Jie Shen, "Efficient Stochastic Galerkin Methods For Random Diffusion Equations", J. Comput. Phys., 228:266-281, 2009.

[5]. Jie Shen and Haijun Yu, "On the approximation of the Fokker-Planck Equation of FENE dumbbell model, I: a new weighted formulation and an optimal spectral-Galerkin algorithm in 2-D", SIAM J. Numer. Anal., in revision.
[6]. Jie Shen and Haijun Yu, "On the approximation of the Fokker-Planck Equation of FENE dumbbell model, II: three dimensional case", preprint.
[7]. G. Wu, G. Lu, C.J. Garcia-Cervera, and W. E, "Density-gradient-corrected embedded atom method", Phys. Rev. B, 79, 035124 (2009).
[8]. C.J. Garcia-Cervera, W. Ren, J. Lu, and W. E, "Sequential multiscale modeling using sparse representation", Comm. Comp. Phys., 4(5), pp. 1025–1033 (2008).
[9]. X. Zhou and W. Ren and W. E, "Adaptive minimum action method for the study of rare events", J. Chem. Phys., 128, 10, 2008.
[10]. X. Wan, X. Zhou and W. E, "Noise-induced transitions in the Kuramoto-Sivashinsky equation", preprint, submitted.
[11]. W. E and X. Zhou, "Noise-induced subcritical instabilities in spatially extended systems", preprint.
[12] L. Lin, J. Lu, L. Ying and W. E, Pole-based approximation of the Fermi-Dirac function, Chin. Ann. Math. 30B, 729, 2009
[13] L. Lin, C. Yang, J. Lu, L. Ying and W. E, A Fast Parallel algorithm for selected inversion of structured sparse matrices with application to 2D electronic structure calculations, accepted by SIAM J. Sci. Comput., 2011
[14] L. Lin, C. Yang, J. Meza, J. Lu, L. Ying and W. E, SelInv – An algorithm for selected inversion of a sparse symmetric matrix, to appear in ACM Trans. Math. Software, 2011
[15] L. Lin, J. Lu, L. Ying, R. Car and W. E, Fast algorithm for extracting the diagonal of the inverse matrix with application to the electronic structure analysis of metallic systems, Commun. Math. Sci. 7, 755, 2009

Honors & Awards Received

Weinan E, Ralph E. Kleinman award, awarded by SIAM, 2009
Weinan E, SIAM Fellow, 2009
Jie Shen, Fulbright Award for McGill University Research Chair, 2008
Jie Shen, Changjiang Chair Professorship, Chinese Ministry of Education, 2010
Lin Lin, Harold Dodds Honorific Fellowship, Princeton University, 2010
Lin Lin, Ray Grimmet Memorial Prize in Computational Physics, Princeton University, 2010
Lin Lin, Best Poster Award, Americal Physical Society March Meeting, 2010.

Transitions

None.